J Sci Food Agric 85:1596–1602 (2005)
DOI: 10.1002/jsfa.2157

Prediction of fat in intact cereal food products using near-infrared reflectance spectroscopy‡

Sandra E Kays,* Douglas D Archibald† and Miryeong Sohn

Quality Assessment Research Unit, Agricultural Research Service, US Department of Agriculture, Russell Research Center, 950 College Station Road, Athens, GA 30605, USA

Abstract: To evaluate the feasibility of an intact product approach to the near-infrared (NIR) determination of fat content, a rapid acquisition spectrometer, with an InGaAs diode-array detector and custom built sampling device, was used to obtain reflectance spectra (1100–1700 nm) of diverse cereal food products. Fat content reference data were obtained gravimetrically by extraction with petroleum ether (AOAC Method 945.16). Using spectral and reference data, partial least-squares regression analysis was applied to calculate a NIR model ($n = 89$) to predict fat in intact cereal products; the model was adequate for rapid screening of samples, predicting the test samples ($n = 44$) with root mean square error of prediction (RMSEP) of 11.8 (range 1.4–204.8) g kg^{-1} and multiple coefficient of determination of 0.98. Repeated repacking and rescanning of the samples did not appreciably improve model performance. The model was expanded to include samples with a broad range of particle sizes and moisture contents without reduction in prediction accuracy for the untreated samples. The regression coefficients for the models calculated indicated that spectral features at 1165, 1215 and 1395 nm, associated with CH stretching in fats, were the most critical for model development.
Published in 2005 for SCI by John Wiley & Sons, Ltd.

Keywords: near infrared; NIR; cereal; food product; fat; rapid analysis; repack

INTRODUCTION

The fat content of food products is an important aspect of consumer food choices and is required for nutrition labeling and monitoring in numerous countries.[1–3] Methods used to measure fat content often require the use of organic solvents and their disposal or repurification.[4,5] Therefore, environmentally benign methods such as near-infrared (NIR) spectroscopy have been developed. Prior NIR calibrations for the successful prediction of fat in mixed cereal foods have been conducted with ground or milled products.[6,7] Ground samples have the advantage that they are more homogeneous in composition and particle size than intact products. Although homogeneity of composition and consistent and small particle size (<1000 microns) are desirable qualities in materials analyzed by spectroscopic techniques,[8] NIR spectroscopy is used successfully to predict protein and moisture in intact wheat grain and has become a standard practice for evaluation of grain quality.[9,10] At the grain elevator or during on-line processing, it is more practical and less time consuming if spectra can be obtained non-destructively on intact grains or products for assessment of composition. With this approach the implementation of NIR technology has been responsible for substantial savings in cost, increased speed of analysis and decreased chemical use.

The current study investigated the feasibility of using NIR reflectance spectroscopy to predict crude fat in a diverse range of intact cereal products, which included ready-to-eat breakfast cereals and cereal-based snacks. A significant variation in particle size is expected in the majority of the samples due to inherent differences in products and incidental breakage and compression during product handling and shipping. Furthermore, variation in moisture content can occur due to differences in ambient relative humidity. Varying conditions of humidity and particle size can affect the performance of NIR methods.[8,11,12] As a consequence, the study included expansion of the NIR calibration to include samples with a wide range of particle sizes and moisture contents for adequate utility in commercial settings. In addition, repetitive packing and scanning of individual samples was employed to ensure acquisition of representative spectra.

* Correspondence to: Sandra E Kays, Quality Assessment Research Unit, Agricultural Research Service, US Department of Agriculture, Russell Research Center, 950 College Station Road, Athens, GA 30605, USA
E-mail: sekays@qaru.ars.usda.gov
†Current address: Crop and Soil Sciences Department, 116 Agricultural Sciences and Industries Building, The Pennsylvania State University, University Park, PA 16802, USA

(*Received 13 September 2004; revised version received 1 November 2004; accepted 10 December 2004*)
Published online 31 March 2005

MATERIALS AND METHODS

Samples, sub-sampling and sample treatments

Cereal food samples were purchased from retail stores and selected so that the model would be robust to the range of cereal products available in the marketplace. Products included breakfast cereals, snack foods, flours, crackers, pastas and baking mixes and comprised a range of grains, including wheat, oat, corn, rice, rye, barley and millet with numerous products having multiple grain types. Many products contained additives, such as dried fruit, nuts, fruit juice, cinnamon, honey, salt, sugar and fat, and were processed by a wide variety of methods including baking, extrusion, milling, frying and air puffing.

Three to six boxes of each product were required to obtain adequate material for sample treatments and spectral and reference analyses. For accurate sub-sampling, the contents of the boxes were mixed and poured onto a rotating platform at a constant rate to give a circular shaped mass. The mass was split along the diameter forming two sub-samples. The sample could be divided further by separating along a diameter at a 90° angle to the first to form four sub-samples. Sub-samples were used for scanning to obtain NIR spectra, for measurement of the parameters, and for relative humidity and particle-size treatments.

To simulate various degrees of breakage, samples were crushed moderately or severely to provide breakage at two levels, giving three particle size treatments (ie unaltered, moderate, severe). This was achieved manually inside heavy polyethylene bags to simulate moderately and severely mishandled products. Particle sizes ranged from tenths of millimeters to centimeters mostly because of the great variety of particle sizes of the original products. To simulate various sample moisture levels outside the existing range, samples were treated by: (1) desiccation using calcium sulfate, and (2) exposure to 60% or >80% relative humidity air in environmental chambers. This produced four moisture levels including the unaltered specimen. Samples with very high fat and sugar content could not be included in the high relative humidity treatments as they became flaccid and, in some cases, partially solubilized by moisture uptake. The particle size and relative humidity treatments expanded the number of specimens from 137 to 361 and the relative humidity treatment increased the moisture content range from the original range of 23–137 g moisture kg^{-1} in the untreated sample set to 22–214 g moisture kg^{-1} in the overall group of treated and untreated samples.

Spectroscopic analysis

Near-infrared spectra were obtained with a Perten Instruments Inc Model DA7000 Spectrometer (Springfield, IL). The instrument has silicon and InGaAs diode arrays and an intense broadband light source, making it possible to measure reflectance from a large area of the sample surface (approximately 10 cm diameter). The diodes were centered at 10 nm intervals but software was used to spline–interpolate spectra to a data interval of 5 nm. The two spectral ranges of the instrument are spliced at 950 nm to cover a range from 400 to 1700 nm. The instrument averages 30 spectra s^{-1} of acquisition time and a spectral scan was defined as the average spectrum generated after 1 s of acquisition. A custom-built hopper with a drop hatch and adjustable sample thickness (2.5–10.0 cm) facilitated loading and unloading the specimens. To ensure obtaining a representative spectrum of each sample and investigate the importance of repacking, each product (including relative humidity and particle-size-treated products) was packed into the hopper against the measuring window eight individual times. For each repack, the instrument was set to collect four scans of the stationary specimen. The instrument reference material is a Spectralon™ disk (Labsphere, Inc, North Sutton, NH), which could be inserted in the custom-built hopper in place of the sample. The reference was scanned after every second sample (ie every two by eight repacks). Most of the partial least-squares (PLS) regression models were generated from the average spectrum for the first repack. Regression models were also developed from the average spectrum of two, four and eight repacks for each sample to compare the results with the first repack.

Analysis of fat content

Crude fat content of the untreated samples was measured by a solvent extraction–gravimetric method using the Soxtec 1040 Extraction System (Foss North America, Inc, Eden Prairie, MN), with petroleum ether as the solvent (AOAC Method 945.16).[4] Dry matter was determined at 105 °C in a forced air oven (AOAC Method 945.14),[13] and crude fat expressed on a dry weight basis.

Calibration development and validation procedure

The original untreated samples ($n = 137$) were divided into calibration ($n = 92$) and validation ($n = 45$) data sets by ranking the samples in increasing order of fat content and assigning each third sample to the validation data set. The moisture and particle size expanded data set ($n = 361$) was divided into calibration ($n = 243$) and validation ($n = 118$) data sets by assigning the treated samples to the same set as their untreated counterparts. Data processing was performed with the Unscrambler v 8.2 software (Camo Inc, Corvallis, OR). Initial analysis indicated that the 1100–1700-nm spectral range was optimal for the study, partly because of the interference from color at shorter wavelengths. Partial least-squares regression (PLS)[14,15] was used to develop calibrations. The use of Martens' Uncertainty regression[16,17] did not afford any advantage in model development for this data set. Spectra for each calibration data set were preprocessed with: (a) multiplicative scatter correction,[18] to partially correct for baseline differences; and (b) second-derivative processing calculated by the Savitzky–Golay

convolution method using a cubic polynomial fit and a window width of seven.[19,20]

Random cross-validation was used to evaluate the models.[15] Pre-processing was the optimum required to improve the root mean square error of cross-validation (RMSECV) of predicted versus reference data for the PLS regression model compared with PLS without preprocessing (or spectra processed with other combinations of scatter correction and derivative). The optimum number of PLS factors used to predict fat content was determined by cross-validation. Performance of each PLS model was first reported as RMSECV, multiple coefficient of determination (R^2) and bias of the calibration. For outlier detection, the x-variables T^2 Hotelling test and y-variables residuals were examined.

Using the model to predict the independent validation samples also tested performance of each PLS model. Performance was reported as the root mean squared error of prediction (RMSEP), which is corrected for bias, the coefficient of determination (r^2), and the bias[15,21] of the linear regression of NIR-predicted versus reference values for fat content. The ratio of deviation to performance, or RPD, is the ratio of the standard deviation of the reference values to the RMSEP and provides a method-independent standardization of the RMSEP.[22] Correlation methods with RPD values of 3.1–4.9 are, in general, considered adequate for screening purposes and values of 5.0–6.4 are adequate for quality control.[22]

Additional models were developed with the original untreated samples ($n = 89$) after two, four and eight repacks using identical processing to the above models. The original calibration and validation samples from the first repack were then combined and a model developed using the full set of untreated samples ($n = 137$). Likewise, a calibration was developed using the full set of untreated and treated samples ($n = 361$). The calibrations with the combined data sets were developed with identical processing to the above models.

RESULTS

Spectral characteristics

The spectra obtained were typical of cereal food products (Fig 1). Products with higher fat content had sharper peaks at 1212 nm, the second overtone region for CH stretching vibrations. There was a broad overall peak around 1450 nm, which is a first overtone region for OH groups. Sharp peaks were observed, within this broader peak, in high sugar/sugar-coated samples at 1430–1440 nm, corresponding with the first overtone for OH groups in carbohydrate.[8,23]

Reference analysis of fat content

Values for crude fat content measured by AOAC Method 945.16[4] ranged from 0.2 to 247 g fat kg^{-1} and the standard error of the laboratory determinations was 2.0 g fat kg^{-1}.[23] Values for the range, mean and standard deviation of the samples in the calibration and validation data sets, and the combined data sets (after the removal of outliers) are presented in Table 1. The ranges of fat content for particle size treated

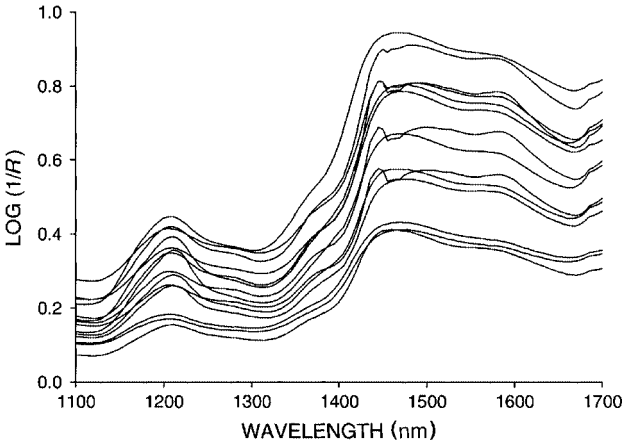

Figure 1. Representative NIR spectra [log (1/R)] of untreated, intact cereal food products. Samples with high fat content have sharper peaks at 1212 nm and samples with high crystalline sugar content have characteristic peaks at 1430–1440 nm.

Table 1. Distribution of fat content (g fat kg^{-1}) for cereal product samples used in the calibration and validation data sets[a]

Data set	*n*	Range	Mean	SD
Original calibration set (untreated samples only)	89	0.2–247.0	41.6	57.9
Original validation set (untreated samples only)	44	1.4–204.8	40.5	54.1
Expanded calibration set (untreated, RH and PS treated samples)	241	0.2–247.0	23.8	36.0
Expanded validation set (untreated, RH and PS treated samples)	117	1.4–204	24.3	35.9
Relative humidity treated samples	25	1.4–37.9	17.6	12.8
Particle size treated samples	26	1.4–37.9	14.0	10.2
Relative humidity and particle size treated samples	22	1.4–37.9	14.6	10.4
Original data set (combined untreated samples)	134	0.2–247.0	41.6	56.3
Expanded data set (combined untreated, RH and PS treated samples)	358	0.2–247.0	24.6	37.0

[a] n = number of samples; SD = standard deviation; RH = relative humidity; PS = particle size.

and relative humidity treated samples are narrower as low fat samples (0–38 g kg^{-1}) were used for these treatments.

Calibration for fat content

Cross-validation and validation statistics for the models developed with the original and expanded data sets are given in Table 2. Overall performance of the original model when predicting the independent validation samples was found to be adequate for screening purposes with RPD values of 4.58. Three outliers, one spectral and two residual, were identified in the original calibration sample set and removed and one residual outlier was identified and removed from the original validation sample set. The same untreated samples were residual outliers in the expanded calibration and validation data sets and were removed; however, the spectral outlier that occurred in the original calibration data set was not an outlier in the expanded calibration data set. The residual outlier in the validation data sets and one of the two residual outliers in the calibration data sets had a NIR predicted value substantially higher than the reference value.

When the validation samples were separated into the treatment types, it was observed that the expanded model predicted the original untreated validation samples with the same accuracy as the original model (Table 3). The expanded calibration predicted the fat content of the relative humidity- and particle-size-treated validation samples with better accuracy than the original model having RMSEP of 8.2–9.0 g kg^{-1}, compared with predictions by the original model, having RMSEP of 11.1–12.2 g kg^{-1} (Table 3). The greatest improvement was seen for the samples with both relative humidity and particle-size treatments.

Averaging two, four or eight repacks did not improve the accuracy of the original model (Table 4). However, the cross-validation performance of the models was improved when calibration and validation samples were combined to develop models with a larger number of samples (Table 5).

Regression coefficients and variation in fat concentration

The regression coefficients show that analytically useful absorptions for the original PLS model ($n = 89$) are at 1165, 1215 and 1395 nm, which are all bands associated with absorption by CH groups in lipids (Fig 2A).[8,24] They are the CH second overtone region at 1165 and 1215 nm and CH combinations region at

Table 3. NIR prediction of fat (g kg^{-1}) by PLS regression models in untreated and treated validation samples[a]

	Validation samples			
Model	Treatment	*n*	RMSEP	r^2
Original ($n = 89$)	None	44	11.8	0.98
	RH	25	11.1	0.70
	PS	26	11.3	0.74
	RH + PS	22	12.2	0.61
Expanded ($n = 241$)	None	44	11.0	0.98
	RH	25	9.0	0.78
	PS	26	9.0	0.79
	RH + PS	22	8.2	0.75

[a] n = number of samples; RMSEP = root mean square error of performance; r^2 = coefficient of determination; RH = relative humidity treated; PS = particle size treated.

Table 4. Calibration and validation statistics of PLS models to predict fat (g kg^{-1}) in the original cereal food products when multiple repacks are used to obtain NIR spectra[a]

	Number of repacks			
Statistics	1	2	4	8
Calibration				
RMSECV	15.8	16.8	15.3	14.8
R^2	0.96	0.96	0.96	0.97
Bias	−0.1	<−0.1	−0.5	−0.4
Slope	0.93	0.93	0.94	0.93
Factors	5	5	5	5
n	89	89	89	89
Validation				
RMSEP	11.8	12.7	11.8	11.2
r^2	0.98	0.98	0.98	0.98
Bias	−0.4	0.2	1.6	1.2
Slope	1.01	1.03	1.01	1.00
n	44	44	44	44

[a] n = number of samples; R^2 = multiple coefficient of determination; RMSECV = root mean square error of cross validation; RMSEP = root mean square error of performance; r^2 = coefficient of determination.

Table 5. Cross-validation statistics of NIR models for prediction of fat (g kg^{-1}) in intact cereal products using combined data sets[a]

Model	*n*	Factors	RMSECV	R^2
Original (combined samples)	133	5	14.0	0.97
Expanded (combined samples)	358	5	12.6	0.94

[a] n = number of samples; R^2 = coefficient of determination; RMSECV = root mean square error of cross-validation.

Table 2. Calibration and validation statistics for PLS regression models for the NIR prediction of fat (g kg^{-1}) in intact cereal products[a]

	Calibration				Validation					
Model	*n*	Factors	RMSECV	R^2	*n*	RMSEP	r^2	Bias	Slope	RPD
Original	89	5	15.8	0.96	44	11.8	0.98	−0.4	1.01	4.58
Expanded	241	5	13.5	0.93	117	11.2	0.98	−1.9	0.95	3.20

[a] n = number of samples; R^2 = multiple coefficient of determination; RMSECV = root mean square error of cross validation; RMSEP = root mean square error of performance; r^2 = coefficient of determination; RPD = standard deviation of the reference method/RMSEP.

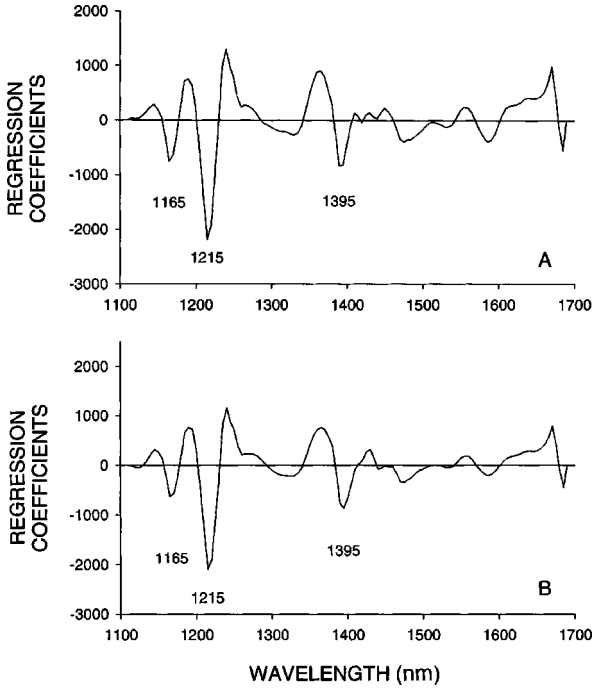

Figure 2. Regression coefficients for the PLS regression model developed with (A) the original, untreated, intact cereal products ($n = 89$) and (B) the expanded data set containing the original untreated cereal products and the relative humidity treated and particle size treated cereal products ($n = 241$).

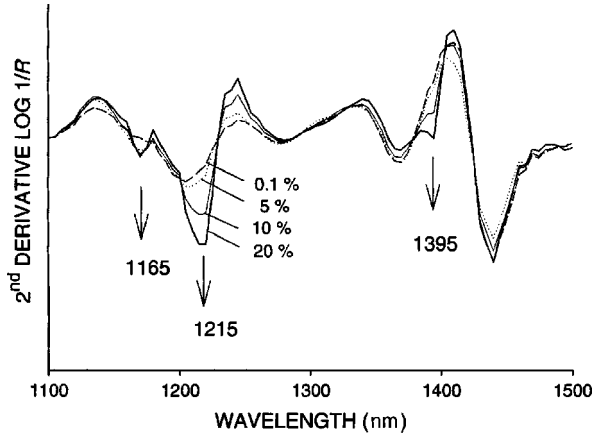

Figure 3. Second derivative NIR spectra of cereal food products containing a broad range in fat content (1, 50, 100, 200 g fat kg^{-1}).

1395 nm. The PLS model with the expanded data set also has major variation at 1165, 1215 and 1395 nm (Fig 2B). Minor variation was also seen in the 1365 and 1480 nm regions. These regions may be associated with CH groups in carbohydrate at 1365 nm and NH groups in protein at 1480 nm. Water does not seem to be involved in either of the two models. Similar observations were made from the regression coefficients of the PLS models developed with the full data sets.

When the second derivative spectra of samples with varying fat contents (1, 50, 100 and 200 g kg^{-1}) are compared, systematic changes in absorbance are observed at 1165, 1215 and 1395 nm (Fig 3). These areas correspond with the areas of greatest influence in the regression coefficients for the PLS models (Fig 2).

DISCUSSION

Traditional measurements of fat by AOAC or AACC methods are very time consuming and often involve the use of organic solvents with the generation of solvent waste.[4,5] Spectroscopic methods, such as NIR spectroscopy, are rapid and eliminate the need for chemicals and chemical waste disposal; therefore NIR spectroscopy was investigated as a possible method of screening products on-line for fat content. NIR spectroscopy has previously, been used for prediction of fat in ground cereal products, with residual standard errors of performance of 9.6–11.0 g kg^{-1} and coefficients of determination (r^2) of 0.98–0.99.[6,7] The current study attempted to develop a model to predict fat content in intact rather than ground cereal products with the aim of eliminating sample preparation time to enable more efficient screening by a manufacturer or food processor. To facilitate this, the instrument used was a rapid-scanning diode-array spectrometer fitted with a custom-made hopper to enable rapid delivery of the sample for scanning. It was found that, using this technique to obtain NIR spectra and using PLS regression, a model could be developed having a root mean square standard error of performance of 11.8 g kg^{-1} (range 1–205 g kg^{-1}) and r^2 value of 0.98. On the basis of these statistics and an RPD of 4.6, the method is suitable for screening purposes and has utility for monitoring fat content in large quantities of cereal products very rapidly without sample preparation. More accurate reference analysis can then be used for samples not in compliance with labelling or not agreeing with expected composition.

A prior model constructed for prediction of total dietary fiber in intact cereal products required 6–7 repacks of the sample to be averaged to obtain optimum model performance.[12] This may be attributed to heterogeneity of the samples due to large differences in particle sizes within and between products and the variation in shapes of individual sample units. In the current study, the averaging of additional repacks did not improve performance of the model to predict fat content. This may be because the signal for fat is stronger than that for dietary fiber and little interference is coming from other constituents. In addition, fat which is added during processing may be more evenly distributed than dietary fiber in the products.

Increasing the scope of the NIR model to include moisture and particle size variation did not greatly affect the accuracy of the model for prediction of fat in the untreated samples scanned 'as is' from the product packaging. The expanded model did have improved prediction errors and r^2 values, compared with the original model, for samples treated with different relative humidity environments. However, the expanded model has not been tested for treated samples with >38 g fat kg^{-1} and further research would be required to test such samples. The limited fat range for treated samples in the expanded model may be responsible for the lower r^2 values observed in the

prediction of these samples, compared with that for the untreated samples. Increasing the variation of particle sizes in the model resulted in improved prediction of crushed or broken samples, indicating that the inherent variation in particle sizes and structures of the products (ranging from flour to breakfast cereals and crackers) was not sufficient to represent additional crushing or breaking. There are several ways that broken or crushed products might be spectrally different from the intact products: (1) by exposure of the interior of products, (2) by the generation of fine articles, and (3) by changing the overall light scattering of the cereal food product. An expanded model of the kind developed here would be useful in circumstances where increased handling of products may cause crushing or breakage and in circumstances where ambient relative humidity is expected to vary. The greatest improvement in prediction results was for the samples with both moisture and particle-size treatments.

Three residual outliers occurred in the data set (one calibration and one validation sample) and the NIR models substantially over-predicted two of them compared with the reference method. In some foods heat treatment during processing can result in alteration of lipids such that they are not readily extracted by solvent alone;[25] therefore, total fat content might have been underestimated by the reference method in two outlier samples.

In conclusion, a near-infrared reflectance model was developed for the prediction of fat in intact cereal products, which is sufficiently accurate for screening samples. Screening can be accomplished rapidly without the need for repacking the sample and without the use of chemicals and the need to dispose of them. The NIR model was expanded to include samples with a wider range in moisture content and particle sizes without loss in prediction accuracy for the original, untreated samples. Regression coefficients suggest that the spectral features that are most important for the models arise from CH stretching in fats with little interference from other components found in cereal food products.

ACKNOWLEDGEMENTS

The authors acknowledge the late Richard B Leffler for fabrication of the custom sampling device for the Perten DA 7000, the technical assistance of Audrey Loper and Meredith Mitchell, and Perten Instruments Inc for the loan of the Perten DA 7000 Spectrometer.

REFERENCES

1 United States Code of Federal Regulations, Title 21, Part 101, Section 9 (2003).
2 Food Standards Agency, Food Law Guide 2004, www.foodstandards.gov.uk (2004).
3 Canadian Food Inspection Agency, 2003 Guide to Food Labelling and Advertizing, www.inspection.gc.ca/english/fssa/labeti/guide (2003).
4 Association of Official Analytical Chemists, Method 945.16, Oil in cereal adjuncts: petroleum ether extraction method, in *AOAC Official Methods of Analysis*, 16th edn. AOAC, Arlington, VA (1995).
5 Association of Official Analytical Chemists, Method 996.01, Fat (total, saturated, unsaturated, and monounsaturated) in cereal products, in *AOAC Official Methods of Analysis*, 17th edn. AOAC, Arlington, VA (1998).
6 Osborne BG, Determination of fat in a variety of cereal foods using NIR spectroscopy, in *Analytical Applications of Spectroscopy*, ed by Crease IS and Davies AMC. Royal Society of Chemistry, London, pp 68–71 (1988).
7 Kays SE, Barton FE II and Windham WR, Analysis of nutritional components in cereal foods with near infrared reflectance spectroscopy, in *Near Infrared Spectroscopy: Proceedings of the 9th International Conference*, ed by Davies AMC and Giangiacomo R. NIR Publications, Chichester, pp 813–816 (2000).
8 Osborne BG, Fearn T and Hindle PH, *Practical NIR Spectroscopy with Applications in the Food and Beverage Industry*, 2nd edn. Longman Scientific and Technical, Harlow, UK (1993).
9 Williams PC and Sobering DC, Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J Near Infrared Spectrosc* **1**:25–32 (1993).
10 Blanco M and Villaroya I, NIR spectroscopy: a rapid response analytical tool. *Trends Anal Chem* **21**:240–250 (2002).
11 Windham WR, Kays SE and Barton FE II, Effect of cereal product residual moisture content on total dietary fiber determined by near-infrared reflectance spectroscopy. *J Agric Food Chem* **45**:140–144 (1997).
12 Archibald DD and Kays SE, Determination of total dietary fiber of intact cereal food products by near-infrared reflectance. *J Agric Food Chem* **48**:4477–4486 (2000).
13 Association of Official Analytical Chemists, Method 945.15, Moisture in cereal adjuncts: air oven method, in *AOAC Official Methods of Analysis*, 15th edn. AOAC, Arlington, VA (1990).
14 Martens M and Martens H, Partial least squares regression, in *Statistical Procedures in Food Research*, ed by Piggott, R. Elsevier Applied Sciences, London, pp 292–360 (1986).
15 Martens H and Naes T, Assessment, validation and choice of calibration method, in *Multivariate Calibration*. John Wiley and Son, New York, pp 237–266 (1989).
16 Martens H and Martens M, Modified Jack-knife estimation of parameter uncertainty in bilinear modeling by partial least squares regression (PLSR). *Food Quality and Preference* **11**:5–16 (2000).
17 Westad F and Martens H, Variable selection in near infrared spectroscopy based on significance testing in partial least squares regression. *J Near Infrared Spectrosc* **8**:117–124 (2000).
18 Martens H, Jensen SA and Geladi P, Multivariate linearity transformation for near-infrared reflectance spectrometry, in *Proceedings of the Nordic Symposium on Applied Statistics*, Stokkand Forlag, Stavanger, Norway, pp 205–234 (1983).
19 Savitzky A and Golay MJE, Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* **36**:1627–1639 (1964).
20 Madden HH, Comments on Savitzky–Golay convolution method for least-squares fit smoothing and differentiation of digital data. *Anal Chem* **50**:1383–1386 (1978).
21 Hruschka WR, Data analysis: wavelength selection methods, in *Near-infrared Technology in the Agricultural and Food Industries*, ed by Williams PC and Norris KH. American Association of Cereal Chemists, St Paul, MN, pp 53–54 (1987).
22 Williams PC, Implementation of near-infrared technology, in *Near-infrared Technology in the Agricultural and Food Industries*, 2nd edn, ed by Williams PC and Norris KH. American Association of Cereal Chemists, St Paul, MN, p 165 (2001).

23 ASTM, Standard Practices for Infrared, Multivariate, Quantitative Analysis. Designation E 1655–94$^{\varepsilon}$, American Society for Testing and Materials, West Conshohoken, PA (1995).
24 Murray I and Williams PC, Chemical principles of near-infrared technology, in *Near-Infrared Technology in the Agricultural and Food Industries*, ed by Williams PC and Norris KH. American Association of Cereal Chemists, St Paul, MN, pp 17–34 (1987).
25 Zou W, Lusk C, Messer D and Lane R, Fat contents of cereal foods: comparison of classical with recently developed extraction techniques. *J AOAC Int* **82**:141–150 (1999).